A platform to support Civil Protection applications on the GRID

António Pina, Bruno Oliveira, Joel Puga, A. Esteves, Alberto Proença

Departamento de Informática, Universidade do Minho
{pina, boliveira, esteves, joel, aproenca}@di.uminho.pt

Abstract. The CROSS-Fire is a Portuguese NGI funded project focusing on the development of a grid-based risk management decision support system for the Civil Protection (CP), using forest fires as the main case study. The project defines a general approach for the development of a CP application by defining an architecture that integrates three main layers: the CROSS-Fire Platform and two external infrastructures: a Spatial Data Infrastructure (SDI) and the GRID.

The CROSS-Fire Platform is defined and implemented as a set of WPS algorithms dealing with most of the functionalities of its three components: Business Logic, Grid Services and Geospatial services.

The present work stresses the relevance of standards adoption: OGC-WS WCS/WFS/WMS/WPS, to exploit/enable geospatial services for data access processing, and OGC-SWE SOS to address other CP data sources, such as meteorological station networks (MSN) or satellites. The adoption of a Web Services (WS) approach allows integrating easily with existing systems typically based on WS technologies.

We also present CFS, a grid user interface SDI based client, compliant with OGC and EU INSPIRE directives which allows decision makers to access the spatial data infrastructure, to launch simulations on the grid and visualize the fire propagation simulations.

Keywords: Civil Protection, Fire Simulation, OGC-WS , OGC-SWE

1 Introduction

The CROSS-Fire is a Portuguese NGI funded Project, leaded by Universidade do Minho (UMinho), focusing on the development of a grid-based risk management decision support system for the Civil Protection (CP), using forest fires as the main case study.

In 2007 UMinho became a member of the EGEE South West federation production infrastructure, advantages of the overall resources available and the facilities of the EGGE infrastructure, including data management, storage and computing. Later, a second site - UMINHO-CP - was installed and configured to support CYCLOPS and EELA VOs, with the objective to offer its researchers a Grid infrastructure well suited to support the requirements of the different types of scientific applications.

1.1 FireStation on the GRID

The base of the project is FireStation (FS), a pre-existent CP standalone application running in a Desktop, developed by the team of Universidade de Coimbra [4]. FS integrates an external module for wind field calculation as well as a FireWeather Index (FWI) module, from the Canadian system to compute the daily fire danger rating. The adopted resolution for the topography terrain and the choice of the wind field generator – Nuatmos, an analytic model developed by Ross et al., 1988 or CANYON that uses a Navier-Stokes solver for a 3D generalized coordinate system [5] – is of the great importance for: i) the quality of the estimated fire spread, ii) the overall computation time and in the amount of input/output data required/produced.

Initial work focused on (i) the parallelisation of the fire propagation execution model and (ii) the integration on the EGEE infrastructure to support higher processing/storage capabilities, to improve I/O data resolution and to allow faster multi-simulation execution and wider simulation areas.

The first FS gLite approach accesses input data stored in a SE, terrain topography and wind field using proprietary formats, generating a text base output that represents the total number of burned cells, the rate of spread, flame height intensity, fire line intensity etc The fire spread simulation takes places on the basis of static information concerning the vegetation cover and the topography of the terrain combined with information for the wind field provided by the wind simulator module. The computation is based on dynamic data for the wind speed and direction manually collected from several meteorological wind stations positioned in place or near the simulation area.

We start by developing a parallel version of FS propagation model using the Message Passing Interface (MPI) also taking advantage of MPI parallel I/O [2] facilities to reduce the overall time needed to process the huge volumes of data. Next, we to extend the P-FS version of the simulator with several currently available Grid tools and services, such as: the LFC catalogue to easy the location, access and transfer of data; the AMGA [9] metadata catalogue to create a database of inputs/outputs of previous executions; and finally the WatchDog [1] tool, to monitor and to provide data for the interactive control of the running simulations.

Both the parallel and the Grid versions were implemented successfully showing that Grid can give a good response to a parallel application, with large input/output demands, allowing for soft real time monitoring and for keeping track of the already computed simulations, through a Grid oriented large meta database.

2 CROSS-Fire

In complex CP applications, the passage from the desktop to the GRID, especially during emergency situations, implies not only the scale from a standalone system to a distributed system, but we must also take into account the need to interact with sources of information that may have strict data policy and security requirements.

In addition, great benefits could be granted to the application by implementing a layer of standard geospatial web-services to access static and dynamic information with specific characteristics, as required by the execution models of FS. The enhancement may opens FS to new data and service providers making possible to conceive more complex scenarios built as a workflows of basic services.

2.1 Architecture

The proposed architecture integrates three main components: the CROSS-Fire Platform and two external infrastructures: a Spatial Data Infrastructure (SDI) platform and the GRID.

The platform is defined and implemented as a set of WPS algorithms dealing with most of the functionalities of its three components: Business Logic, Grid Services and Geospatial services. The Business Logic is an abstract layer containing the application specific semantic that allows users to access the GRID and SDI infrastructures, configured to handle the algorithms that provide all the functionality of FS, namely, forest fire propagation, wind field calculation and FWI modules.

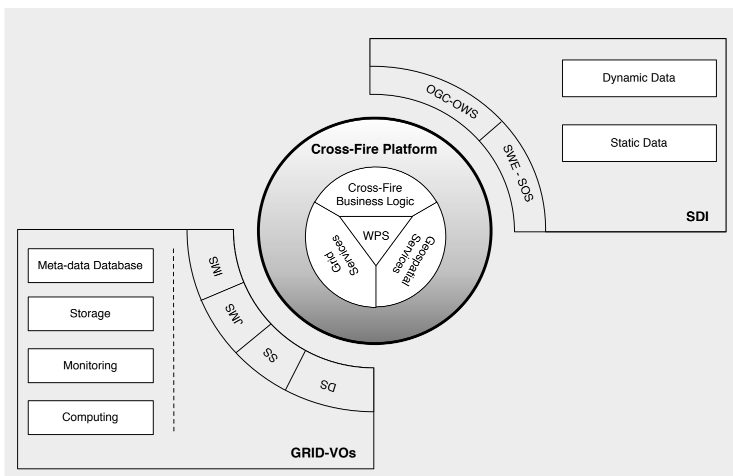

Fig. 1. Bridging the gap between applications, geospatial data and the Grid

The adoption of a web services (WS) approach allows integrating easily CROSS-Fire with existing systems based on WS technologies. The Grid layers make possible the coordinated sharing of basic resources for heavy computation and huge datasets storage, while the SDI layer is used to access and share geo-spatial information stored in standard formats through standard interfaces.

The platform was designed to supports other fire spread models application but it has the potential to support other natural risk assessment and management, such as flash floods, volcanic eruptions and tsunamis.

The GRID infrastructure in Figure 1 follows the Service Oriented Architecture (SOA) paradigm, simplifying interoperability among Grid services and allowing easier compliance with upcoming standards.

The Grid services can be in general divided in the following groups: i) Information & Monitoring Services (IMS), which provides mechanisms to collect and publish information about the state of grid services and resources, as well as to discover them; ii) Job Management Services (JMS), which concerns the execution and control of computational jobs for their whole lifetime through the Computing Element (CE) that provides an interface to access and manage a batch queue of a cluster according to user's job requirements and defined VO and resource-level policies; iii) Security Services (SS) which concerns the Grid Security Infrastructure (GSI) and the gLite Virtual Organisation Membership Service (VOMS), an attribute authority that plays an important role by allowing fine-grained access control to the resources; iv) Data Services (DS) which concerns the access, transfer and cataloguing of Data at the file level. The Storage Element (SE) that provides an interface to one or more storage resources and the LCG File Catalogue (LFC) service that keeps track of the locations of the files and of their replicas distributed in the grid, as well as relevant metadata.

INSPIRE [3] is an EC initiative to build an European SDI infrastructure oriented to geographic data beyond national boundaries that aims to provide relevant, harmonized and quality geographic information, to handle issues concerning data formats, metadata, and network services and to enable a basic set of standard geospatial services.

Currently European/International projects are designing, developing, implementing and using SDI(s) as information infrastructures oriented to geographic data.

3 Open Geospatial Consortium

The OGC-Web Services [7] is being adopted worldwide as the technology to support the development of complex distributed applications over grid platforms, and to deal with data from many different sources, including live sensors.

OGC services provide the conceptual tools for modelling the geospatial information that is shared in CP applications during its entire lifecycle implemented on top of open Internet standards. From the CROSS-Fire viewpoint the standards that follows are the most interesting:

i) Web Coverage Service (WCS) - standardizes access to geospatial data as coverages" representing space-varying phenomena, usually encoded in a binary format and offered by a server for client-side rendering;
ii) Web Map Service (WMS) - standardizes the display of registered and superimposed map like views of information that come simultaneously from multiple remote and heterogeneous sources;
iii) Web Features Service (WFS), standardizes the retrieval and update of digital representations of real-world entities referenced to the Earths surface;
iv) Web Processing Service (WPS), define basic request-response interaction for remote execution of a service, which can include any algorithm, calculation, or model that operates on spatially referenced data, clearly serving as an interface to a wide range of computing resources, from local servers to high-end computational grid resources;

3.1 Sensor Web Enablement

The OGC Sensor Web Enablement (SWE) initiative makes all types of sensors, instruments, and imaging devices accessible and, where applicable, controllable via the Web, through a common interface and encodings.

The SWE/OGC standards are being used to provide CROSS-Fire wind field and FWI execution models with dynamic data, coming from satellites such as Terra/Aqua and weather station O&M such as DAVIS Vantage Pro2 wireless and the associated database.

3.2 OGC-SDI Integration

Figure 2, show the exploitation of Geospatial Services using a standard-based SDI layer to provide FS with both static and dynamic data, by using OGC standards: OGC-WCS/WFS/WPS to exploit/enable geospatial services for data access and processing, and SWE/OGC to address other CP data sources, such as meteorological stations data or satellite images.

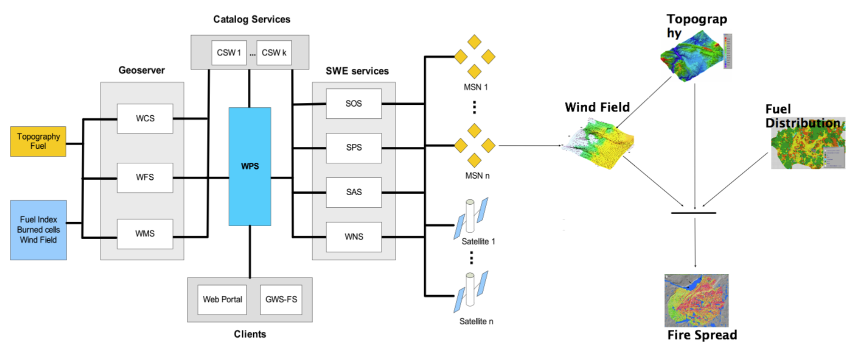

Fig. 2. OGC-SDI Integration

4 WPS algorithms

Figure 3 is another view of the Cross-Fire platform, the SDI and Grid infrastructures on the top, the WPS layer one the centre and a client layer at the button; the white rectangles details the algorithms/services relevant to the presentation of the platform.

The WPS layer is being developed on top of WPS 52 North. In the following sections we will describe some of its algorithms.

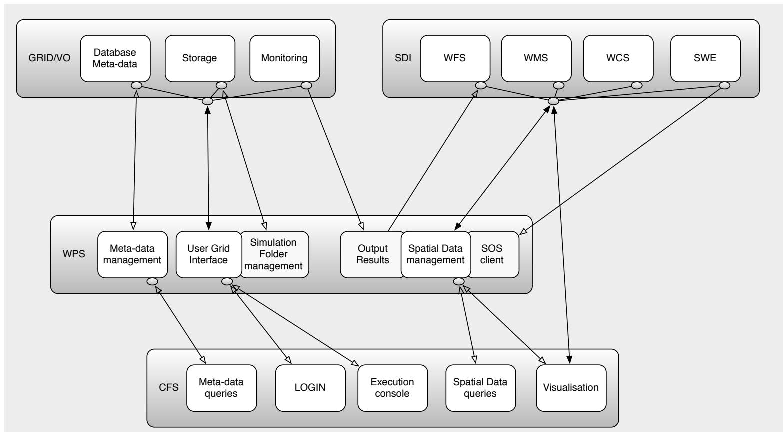

Fig. 3. The CROSS-Fire platform

4.1 Meta-data management

In CROSS-Fire one expects not only to be able to run simulations as a component of the risk management decision system but also to access and manipulate the results produced in the past. To accomplish these requirements the platform must be able to register in a convenient way all input data used in the simulation, as well as outputs.

To achieve those goals, all the simulations input/output data, along with the meta-data that characterize the executions, such as the date or the execution time or the resources used, is recorded and accessed through a specially designed database, supported by a database server with GIS extension.

In order to have this data available on the GRID and, at the same time, to have access to the replication, redundancy and security facilities offered by the GRID, the database is interfaced via AMGA, which allows a uniform view over a set of databases under its management, regardless of the underlying database server architecture.

4.2 User Grid Interface

This algorithm is used whenever a user needs to run a simulation or some data is made available on the GRID. User interacts with this algorithm on the assumption that he already possesses a valid proxy to the GRI/SDI infrastructures. Functionalities include the management of GRID jobs, along with creation of the JDL files, the job submission and the execution control and status.

To launch a simulation, the user identifies a set of geo-referenced files within the SDI, along with a set of control parameters, such as ignition points, time and date and both statically or dynamically meteorological conditions.

To ease the implementation of this algorithm we are currently profiting from the possibilities of the GANGA, which is a tool for computational-task management and easy access to Grid resources [6].

4.3 Simulation folder management

A simulation folder is the physical representation of the meta-data available in the AMGA database. It is a hierarchical data structure created on the GRID where all input/output data needed/produced by a simulation resides.

The objective is not only to create a database of past simulations to post-mortem analysis but also to accelerate the running of the simulation whenever a new job needs data already available as that was used or produced by previous executions.

The simulation folder management algorithm may also be used to allow users to retrieve from the grid the raw data, by issuing a request to the WPS layer.

5 CFS Client

The desktop version of FS offers a nice and user-friendly CAD environment that we decide to augment with facilities to handle geo-referenced data and to interface with both the GRID and the SDI infrastructures.

The console FireStation (CFS) is based on gvSIG a full feature Open Source GIS desktop solution, funded by EC, which conforms to INSPIRE for managing geospatial information.

5.1 Execution Console

This module is used to establish the connection between clients and the GRID, being responsible by the negotiation with the WPS layer the user proxy generation and delegation, to allow users to log in. It also gather and organizes all the data needed to construct the workflow that will run in the GRID the jobs that represent all the execution models comprised in the overall CROSS-Fire Business Logic.

Any other interactions with GRID as it is the case of the access to the Data Services, for instance, to retrieve a file from a Storage Element is also managed by this module.

5.2 Visualization

To interact with the WPS layer of the CROSS-Fire Platform we extended CFS with a WPS client that is been used for built several other modules, including the one used to visualise fire spread simulations.

This module provides a GUI to discover and obtain the information produced by past-executed simulations and visualizes it as an animation of the fire spread, see in figure 5 a screen shot of an instanced of fire spread simulation. Users may also walk back and forth in time; to review fire spread evolution, moving a slider that represents the overall time of the simulation execution.

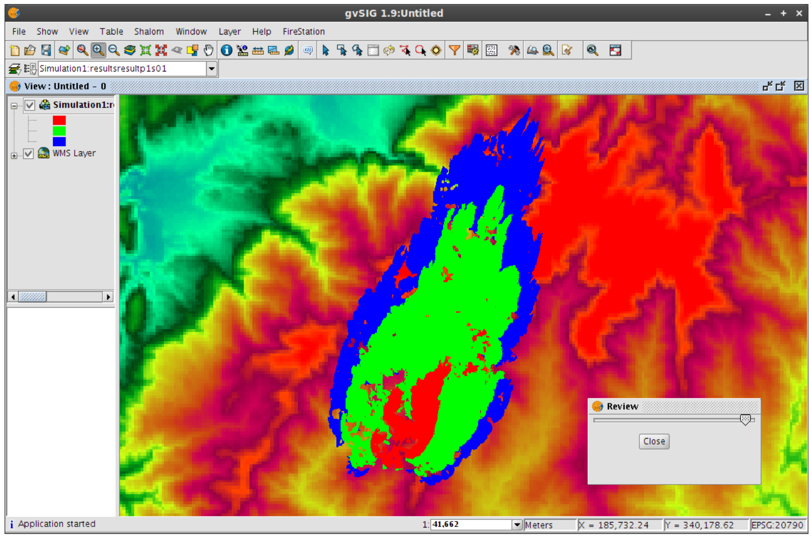

Fig. 4. Visualization with CFS

5.3 Meta-data queries

This module is used to interact with the WPS – Meta-data management – algorithm, by means of temporal or geographic queries, to seek and obtain data or determine the running status of a simulation. This last example is of high importance to the Execution Console module - responsible for the execution of the simulations since it needs to know when new data is made available during the overall execution time.

Users may to explore and use the information that represents previously executed simulations. After submitting a query to the meta-data database to discover past simulations the user may apply the returned information to select the data he wants to review/resubmit. The module may then perform all the necessary tasks, by downloading the corresponding data from the SDI.

5.4 Spatial Data Queries

In order to obtain all the necessary data to run simulations from the SDI, this algorithm interacts with the WPS algorithm Spatial Data Management, allowing users to query the SDI about the availability of data in a certain geographical region.

This module interacts with the user giving then the possibility to choose the kind of data to be obtained terrain, fuel, and meteorological data, among others and to specify parameters, such as geospatial co-ordinates or time intervals. The chosen data is then made available in other layers inside the application.

6 WPS workflow

The WPS layer is the core component of CROSS-Fire architecture. It wraps FS application providing both a standard compliant access interface to the SDI and to the GRID distributed processing capabilities that hides from clients/users the precise syntax of scripting languages and command-line options. It uses GANGA API to work directly with program executables being able to run several independent processes in the background, but having no dependency handling facility, both over the compute nodes of a cluster or of the Grid.

Underlying almost all requests there is a custom XML data structure that serves many purposes, such as, data exchange, query, and process record. The structure matches most of fields of the Schema used by the overall meta-data database.

In what follows we briefly describes the workflows of data acquisition and simulation launching. The depicted figures contemplate both CFS and Grid views; the arrows represented with solid or dotted lines means respectively request and response interactions between boxes.

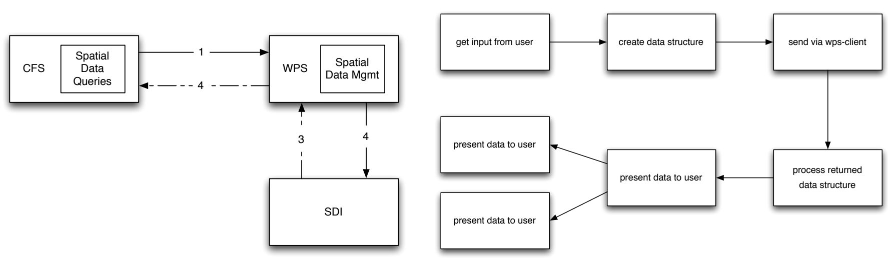

Fig. 5. Data acquisition l) operators view r) CFS view

6.1 Data acquisition

Operator view Assuming the operator logged in the platform through the CFS, he may start an acquisition process using the GUI to create/modify a XML data structure with the information necessary to define the bounding-box corresponding to a certain area of the globe, and sent it (1) as a request to the WPS server -spatial data management- algorithm for querying the SDI for the specific data.

- The WPS contacts the SDI (2) that returns the XML data structure filled (by the GEO Server) with layers that represent terrain, fuel map, fuel description, previously computed DTM's (for wind calculation) or wind-fields, and meteorological stations, whose bounding boxes overlap the requested bounding box - these layers are available though different standards: WCS, WMS or WFS, according to its characteristics.

– Once CFS receives the returned XML data structure (4) the operator can select which layers to use - one for terrain, one for fuel map, one for fuel description, one or more meteorological stations, eventually a DTM and/or a wind field.
– Afterwards the operator can select the desired meteorological stations and retrieve the corresponding meteorological data, filtered by date or interval, by invoking the SDI (dynamic data services), or manually enter the data for these same stations.

CFS view Uses the GUI to get input from the operator, to create the XML data structure and to send the request (1 above) to the WPS-server by means of the CFS-WPS client interface. Afterwards the WPS server returns the updated XML structure (4 above). The operator may select separately each desired layer, that are then downloaded from the SDI and rendered in the CFS CAD environment.

6.2 Simulation request

Operator view Once the operator gather all the necessary information and define the ignition points, a simulation may be launched. This request consists of the XML data structure fully filled that is sent to the WPS server – user grid interface – algorithm using a WPS-client request (1). As defined in the meta-data database Schema a simulation comprises one or more executions.

This algorithm has to perform several tasks, which will depend on the kind of simulation call. A complete request entails a new wind field calculation, which, in turn, may need a new DTM calculation, and finally a fire spread calculation. Requests can, however, reuse information from previous simulations/executions. In effect, a new fire simulation execution may reprocess a previously computed DTMs and/or wind field if only the ignition points/barriers/control are changing.

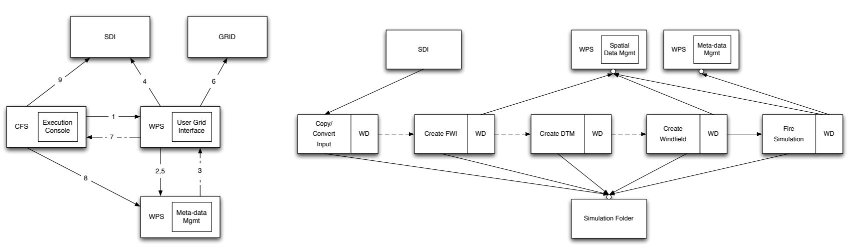

Fig. 6. Lauching the simulation l) Operator view r) Grid View

– Independently from the number of steps, the WPS needs to issue a request to the internal – meta-data management – algorithm (2). This algorithm records the meta-data of the new simulation: simulation, executions, wind fields and DTM's. An identifier is generated if any of these objects – in database terms

– has never been registered before. These identifiers are returned within the XML data structure (3) and are then used when the WPS requests SDI to create new layers (4) to accommodate the results of the simulation. With the information regarding the URLs of the new layers, the revised XML data structure (5) is used to bring the meta-database up-to-date.
- Finally, the described workflow is converted into jobs to be submitted and executed on the grid (6), concluding the simulation request (7).

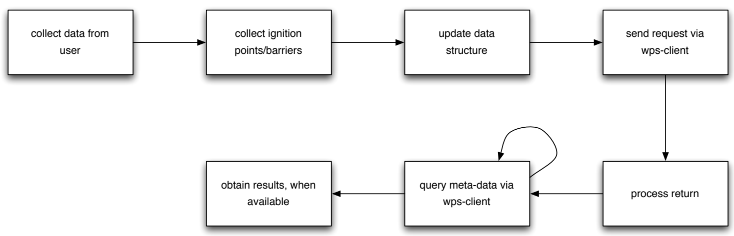

Fig. 7. Lauching the simulation - CFS View

GRID view Next we explain, based in fig. 6 the flux of interactions between the different layers triggered by a clients/operator request of a new simulation/execution.

- As depicted in fig. 6 1) the SDI is the source of the input data needed by the execution models wrapped by CROSS-Fire Business Logic, The initial data (static) includes terrain, fuel, and fuel models accessible in standard formats from the SDI, that needs to be converted to a suitable format before copied to the corresponding Simulation Folder on the SE accessible by GRID/VO. All the boxes in the middle of the figure are divided in two parts: in the left rectangle a specific FS algorithm and in the right rectangle the attached WD script in charge of monitoring the algorithm.

CFS view

- Uses the GUI to collect data corresponding to the ignition points/barriers/-control used to update the XML data structure with the information needed to launch the simulation by sending a request (1 above) to the WPS server and returning (7 above).
- Using the identifiers returned (7 above) the client/operator gets a continuous update of the job(s) submitted the GRID, since CFS does frequents WPS-client requests to the meta-data database (8). Once the jobs are in a running status, CFS inquires the SDI (9) to obtain the execution results, and renders them.

7 Conclusions

In this paper we presented CROSS-Fire - a platform based on WS technologies - that provides both a standard compliant access interface to the SDI and to the GRID distributed processing capabilities. It was designed to support a risk management support system, with real-time or near real-time availability of critical geo-referenced data and settings-based forecasts for fire spreading. The platform may also be used to support new brands of CP applications, such as Flash floods, implemented as a new Application Business Logic component that responds to the specificities of the new application domain.

Current work is centred on the developing of a: 1)parallel implementation of CANYON wind simulation model to reduce the time of overall fire spread simulation, 2) WPS-server X509 based module for a One Sign Authentication Management, to allow a consistent integrating of user access to both the GRID and the SDI and 3)decision-support system based on a web portal where many players can connect, to request services through the core WPS layer.

8 Acknowledgements

This work is supported by the FCT grant GRID/GRI/81795/2006 awarded to the CROSS-Fire project. UMinho is also member of EC FP7 EELA-2 project taking advantage of the prod.vo.eu-eela.eu VO to develop the FS grid prototype. Finally, the authors would like to acknowledge the work of all team members, without whom the project would not be possible: A. Fonseca, B. Gomes, L. Miranda and M. Caldas, and V. Oliveira for keeping the machines up and running smoothly.

References

1. Bruno, R., A watchdog utility to keep track of job execution, INFN, https://grid.ct.infn.it/twiki/bin/view/PI2S2/WatchdogUtility
2. Gropp, W., Thakur, R., and Lusk, E. ,Using MPI-2: Advanced Features of the Message Passing Interface, MIT Press Cambridge, MA, USA, 1999
3. INSPIRE, http://inspire.jrc.ec.europa.eu/
4. Lopes, A., FireStation Users Manual, Universidade de Coimbra, 2000
5. Lopes, A., Cruz, M. , and Viegas, D., FireStation - an integrated software system for the numerical simulation of fire spread on complex topography, Environmental Modelling and Software, 17(3):269285, 2002
6. Mocicki, J.T., Brochu, F., et all, Ganga: A tool for computational-task management and easy access to Grid resources, Computer Physics Communications, Volume 180, Issue 11, November 2009, Pages 2303-2316
7. Open Geospation Consortium, http://www.opengeospatial.org/
8. Pina, A., Oliveira, B., Puga, J., Esteves, A., and Proença, A., An OGC-WS framework to run FireStation on the Grid , 3rd Iberian Grid Infrastructure Conference, Valencia, 2009
9. Santos, N., and Koblitz, B., Metadata services on the Grid, Proceedings of the X International Workshop on Advanced Computing and Analysis Techniques in Physics Research, 2005